Chapitre 3 : Tests d'hypothèses.

1 Introduction aux tests d'hypothèses.

1.1 Un exemple.

Dans le cadre de l'exercice 6, on observe un 10-échantillon $(X_1, \cdots, X_{10})$ de la loi $\mathcal{N}(\mu, \sigma^2)$ et on souhaite savoir si $\mu \le 99\text{kg}$, ou si $\mu > 99\text{kg}$. Pour décider entre ces deux hypothèses, on peut se baser sur l'estimateur $\hat{\mu}_n = \bar{X}_{10}$ de μ. Par exemple, on peut adopter la règle de décision suivante :

- Si $\hat{\mu}_n > 99$, on décide que $\mu > 99$.
- Si $\hat{\mu}_n \le 99$, on décide que $\mu \le 99$.

Calculons les probabilités de se tromper si on adopte cette règle. Une première façon de se tromper est de décider que $\mu > 99$, alors que ce n'est pas le cas. La probabilité d'erreur associée est au pire des cas,

$$\alpha_1 = \sup_{\mu \le 99, \sigma > 0} \mathbb{P}_{\mu,\sigma}\left[\hat{\mu}_n > 99\right] = \sup_{\mu \le 99, \sigma > 0} \mathbb{P}_{\mu,\sigma}\left[\frac{\hat{\mu}_n - \mu}{\frac{\sigma}{\sqrt{n}}} > \frac{99 - \mu}{\frac{\sigma}{\sqrt{n}}}\right]$$

Notons F la fonction de répartition de $\mathcal{N}(0,1)$. Puisque $\hat{\mu}_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$, on obtient

$$\alpha_1 = \sup_{\mu \le 99, \sigma > 0} 1 - F\left(\frac{99 - \mu}{\frac{\sigma}{\sqrt{n}}}\right) = \sup_{x > 0}(1 - F(x)) = 1 - F(0) = 1/2,$$

On a donc au pire des cas (qui correspond à $\mu = 99$), 50 % de chances de se tromper en décidant que $\mu > 99$ si on adopte cette stratégie de décision.

L'autre erreur possible est de décider que $\mu \le 99$, alors que ce n'est pas le cas :

$$\begin{aligned}\alpha_2 &= \sup_{\mu > 99, \sigma > 0} \mathbb{P}_{\mu,\sigma}\left[\hat{\mu}_n \le 99\right] = \sup_{\mu > 99, \sigma > 0} \mathbb{P}_{\mu,\sigma}\left[\frac{\hat{\mu}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \le \frac{99 - \mu}{\frac{\sigma}{\sqrt{n}}}\right] \\ &= \sup_{\mu > 99, \sigma > 0} F\left(\frac{99 - \mu}{\frac{\sigma}{\sqrt{n}}}\right) = \sup_{x < 0} F(x) = F(0) = 1/2.\end{aligned}$$

Dans le pire des cas, on a donc aussi 50 % de chances de se tromper en décidant que $\mu \le 99$. Ainsi, avec la règle de décision que nous avons choisie, nous avons 50 % de chances de nous tromper quelle que soit la conclusion à laquelle on aboutit. Autant jouer à pile ou face ! Comment améliorer notre règle de décision ? Au lieu de comparer $\hat{\mu}_n$ à 99kg, on pourrait le comparer à une autre valeur t en essayant d'ajuster t de façon à réduire les probabilités d'erreur précédentes. Si on reprend les expressions précédentes, on a maintenant

$$\alpha_1 = \sup_{\mu \le 99, \sigma > 0} \mathbb{P}_{\mu,\sigma}\left[\hat{\mu}_n > t\right], \qquad \alpha_2 = \sup_{\mu > 99, \sigma > 0} \mathbb{P}_{\mu,\sigma}\left[\hat{\mu}_n \le t\right].$$

Il est clair sur ces expressions que plus t augmente, plus α_1 devient petit. En revanche, plus t augmente, plus α_2 devient grand. **Les deux erreurs ne peuvent pas en général être petites en même temps.**

La méthode utilisée est alors

- de choisir parmi les deux types d'erreurs possibles, l'erreur que l'on veut absolument contrôler (parce que par exemple, elle a des conséquences plus coûteuses que l'autre) ;
- de lui assigner un niveau α petit ;
- de choisir t en conséquence.

Par exemple, dans le cas de l'exercice 6, on peut vouloir éviter de conclure que le nouveau procédé est meilleur alors qu'il ne l'est pas. L'erreur que l'on veut contrôler est donc

$$\alpha_1 = \sup_{\mu \le 99, \sigma > 0} \mathbb{P}_{\mu,\sigma}\left[\hat{\mu}_n > t\right].$$

En pratique, cela revient à privilégier une hypothèse par rapport à l'autre. On va tester (H_0) : "$\mu \le 99kg$", contre (H_1) : "$\mu > 99kg$". Par **définition de $(\mathbf{H}_0)$** (et donc de (H_1)), l'erreur que l'on contrôle est

$$\boldsymbol{\alpha = P_{(H_0)}(\mathrm{Rejeter}(H_0))}.$$

α s'appelle **le niveau du test**, ou **l'erreur de première espèce.** Dans l'exemple choisi, pour trouver une règle de décision de niveau α, nous allons devoir changer la statistique utilisée. En effet, pour tout $t \in \mathbb{R}$,

$$\begin{aligned}\sup_{\mu\leq 99,\sigma>0} \mathbb{P}_{\mu,\sigma}\left[\hat{\mu}_n > t\right] &= \sup_{\mu\leq 99,\sigma>0} 1 - F\left(\frac{t-\mu}{\frac{\sigma}{\sqrt{n}}}\right) = 1 - \inf_{\mu\leq 99,\sigma>0} F\left(\frac{t-\mu}{\frac{\sigma}{\sqrt{n}}}\right) \\ &= 1 - \inf_{\sigma>0} F\left(\frac{t-99}{\frac{\sigma}{\sqrt{n}}}\right) \geq 1 - F(0) = 1/2,\end{aligned}$$

en prenant $\sigma \to +\infty$ dans l'infimum. Or, considérer le cas $\sigma \to +\infty$ n'est pas forcément pertinent, puisqu'on a une estimation de σ donnée par S_n', où $S_n'^2$ est l'estimateur sans biais de la variance. Aussi, on va plutôt utiliser comme statistique de test $(\hat{\mu}_n - 99)/(S_n'/\sqrt{n})$. La règle de décision va être la suivante :

— Si $(\hat{\mu}_n - 99)/(S_n'/\sqrt{n}) > t$, on rejette (H_0) ;
— Si $(\hat{\mu}_n - 99)/(S_n'/\sqrt{n}) \leq t$, on ne rejette pas (H_0).

t est choisi de telle sorte que

$$\alpha = 5\% = \sup_{\mu\leq 99,\sigma>0} \mathbb{P}_{\mu,\sigma}\left[\frac{\hat{\mu}_n - 99}{\frac{S_n'}{\sqrt{n}}} > t\right] = \sup_{\mu\leq 99,\sigma>0} \mathbb{P}_{\mu,\sigma}\left[\frac{Z + \frac{\mu-99}{\sigma/\sqrt{n}}}{\sqrt{\frac{Y}{n-1}}} > t\right],$$

où $Z = (\hat{\mu}_n - \mu)/(\sigma/\sqrt{n})$ et $Y = (n-1)S_n'^2/\sigma^2$. On peut montrer que Z et Y sont indépendantes, avec $Z \sim \mathcal{N}(0;1)$, et $Y \sim \chi^2_{n-1}$. Par conséquent, le suprémum en μ est atteint en $\mu = 99$, et on a donc

$$\alpha = 5\% = \sup_{\mu\leq 99,\sigma>0} \mathbb{P}_{\mu,\sigma}\left[\frac{Z}{\sqrt{\frac{Y}{n-1}}} > t\right] = \mathbb{P}\left[\frac{Z}{\sqrt{\frac{Y}{n-1}}} > t\right]$$

Par définition de la loi de Student, $Z/\sqrt{Y/(n-1)} \sim St_{n-1}$ et t est donc le quantile d'ordre 95% de la loi St_{n-1} : $t = t_{n-1;0.95}$.

Une fois le niveau α fixé, t est fixé, la régle de décision est fixée, et **on n'a plus aucun moyen de contrôler le deuxième type d'erreur**, à savoir

$$\textbf{erreur de deuxième espèce} = P_{(H_1)}(\textbf{accepter } (H_0)).$$

Dans l'exemple précédent, cette erreur est donnée par la fonction

$$(\mu,\sigma) \in]99;+\infty[\times\mathbb{R}^+ \mapsto \mathbb{P}_{\mu,\sigma}\left[\frac{\hat{\mu}_n - 99}{\frac{S_n'}{\sqrt{n}}} \leq t_{n-1;0.95}\right] = \mathbb{P}\left[\frac{Z + \frac{\mu-99}{\sigma/\sqrt{n}}}{\sqrt{\frac{Y}{n-1}}} \leq t_{n-1;0.95}\right].$$

Cette fonction est décroissante en μ, et lorsque $\mu \to 99$, la valeur de cette erreur tend vers $\mathbb{P}\left[Z/\sqrt{Y/(n-1)} \leq t_{n-1;0.95}\right] = 95\%$. Ainsi, l'erreur de deuxième espèce peut être élevée suivant les valeurs du paramètre. **Il ne faut pas s'en étonner !** Plus μ est proche de 99kg (par valeurs supérieures), et plus il est difficile de voir sur un échantillon si μ est plus grand ou plus petit que 99.

L'absence de contrôle sur l'erreur de deuxième espèce a la conséquence suivante :

la conclusion d'un test n'a valeur de preuve que lorsque cette conclusion est le rejet de (H_0).

En effet, si à l'issue du test, on conclut au rejet de (H_0), on sait qu'on a une probabilité α de se tromper. Si au contraire, on conclut à l'acceptation de (H_0), on a peut-être grande chance de se tromper. Ainsi, il faut plutôt voir l'acceptation de (H_0) comme un non-rejet de (H_0). La démarche est ici très empirique : sur l'expérience que j'ai faite, rien ne permet de dire que (H_0) n'est pas vraie. Comme il est dit dans le livre de Wasserman (All of Statistics : a concise course in Statistical Inference, Springer texts in statistics) :
"Hypothesis testing is like a legal trial. We assume someone is innocent unless the evidence strongly suggests that he is guilty. Similarly, we retain (H_0) unless there is strong evidence to reject (H_0)".

1.2 Définitions

De façon plus générale, on peut formuler les définitions suivantes.

Définition 1. Soit $\mathcal{T}_0 \cup \mathcal{T}_1$ une partition de l'ensemble des paramètres $\mathcal{T}$. Un **test de niveau α de l'hypothèse** (H_0) : **"$\theta \in \mathcal{T}_0$" contre l'hypothèse** (H_1) : **"$\theta \in \mathcal{T}_1$"**, est la donnée d'un sous-ensemble $\mathcal{R}$ de l'ensemble des observations $\mathcal{X}$ ($\mathcal{R}$ ne dépendant pas du paramètre θ) tel que

$$\sup_{\theta \in \mathcal{T}_0} \mathbb{P}_\theta [X \in \mathcal{R}] \leq \alpha. \tag{1}$$

- $\mathcal{R}$ s'appelle la **zone (ou région) de rejet de** (H_0).
- (H_0) s'appelle **l'hypothèse nulle**. C'est l'hypothèse privilégiée de celui qui fait le test, en ce sens qu'il ne veut pas se tromper en la rejetant. (H_1) est **l'hypothèse alternative**
- La **règle de décision** associée à un test de région de rejet $\mathcal{R}$ consiste à décider que $\theta \notin \mathcal{T}_0$ (rejet de l'hypothèse (H_0)) dès que $X \in \mathcal{R}$ (l'observation est dans la région de rejet). Autrement dit, en notant 1 la décision de rejeter (H_0) et 0 la décision contraire, la règle de décision est la fonction $d : \mathcal{X} \to \{0, 1\}$ donnée par :

$$d(X) = \begin{cases} 1 & \text{si } X \in \mathcal{R}, \\ 0 & \text{sinon.} \end{cases}$$

- La **taille du test** est $\sup_{\theta \in \mathcal{T}_0} \mathbb{P}_\theta [X \in \mathcal{R}] = \sup_{\theta \in \mathcal{T}_0} \mathbb{P}_\theta [d(X) = 1]$.
- Lorsque ce maximum est atteint en une seule valeur θ^* de $\mathcal{T}$, cette valeur correspond au **cas le plus difficile**. C'est la valeur de θ pour laquelle il est le plus difficile de prendre une décision si les données en sont issues.
- La **fonction puissance du test** est la fonction $\beta : \theta \in \mathcal{T}_1 \mapsto \mathbb{P}_\theta [X \in \mathcal{R}] = \mathbb{P}_\theta [d(X) = 1]$ (probabilité de rejeter à raison (H_0)).
- L' **erreur de deuxième espèce du test** est la fonction $\theta \in \mathcal{T}_1 \mapsto 1 - \beta(\theta)$.

Remarque 2. Si on a le choix entre plusieurs tests de même niveau, on choisira évidemment celui qui a la plus grande fonction puissance, dans l'hypothèse où ces fonctions peuvent être comparées.

Remarque 3. Lorsque le niveau α change, la zone de rejet, ainsi que la règle de décision changent. Lorsque l'on veut étudier toute une collection de tests de niveau α différents, il est courant d'indexer par α la région de confiance $\mathcal{R}_\alpha$, ainsi que la règle de décision d_α.

1.3 Construction d'un test

La construction d'un test suit alors les étapes suivantes :

1. **Choix de l'hypothèse privilégiée** (H_0). (H_0) doit être choisie en fonction de l'erreur que l'on veut contrôler "$\alpha = P_{(H_0)}$(rejeter H_0)".
2. **Choix du niveau α (petit).** La valeur traditionnelle est $\alpha = 5\%$.
3. **Choix d'une statistique de test.** On choisit une statistique $T = t(X)$ dont le comportement est le plus différent possible entre (H_0) et (H_1). (Par exemple, un estimateur $\hat{\theta}(X)$ du paramètre θ.)
4. **Détermination de la zone de rejet (à des bornes près)** $\mathcal{R}$ en fonction du comportement de la statistique T lorsque θ est très loin de $\mathcal{T}_0$: $\mathcal{R} = \{x \in \mathcal{X} \ : \ t(x) \in \ldots\}$
5. **Détermination des bornes de zone de rejet au niveau α**, en fonction de la loi de la statistique sous (H_0), pour que l'équation (1) soit vérifiée.
6. **Etude de la fonction puissance** (ou ce qui revient au même de l'erreur de deuxième espèce) du test ainsi construit.

Si on a plusieurs tests à sa disposition, une dernière étape consiste à choisir, parmi ces différents tests, celui qui a la plus grande fonction puissance (s'il y en a un).

Remarque 4. On doit donc connaître la loi de la statistique de test sous (H_0) pour pouvoir construire la région de rejet. Cette contrainte impose souvent le choix de l'hypothèse (H_0).

Remarque 5. Les notions de région de confiance et de zone de rejet sont en dualité. On montrera en exercice que pour toute région de confiance $I(X)$ de coefficient de sécurité $1 - \alpha$, on peut définir un test de niveau α de $(H_0) : \theta = \theta_0$ contre $(H_1) : \theta \neq \theta_0$ en choisissant comme région de rejet $\mathcal{R} = \{x \in \mathcal{X}$ tel que $\theta_0 \notin I(x)\}$.

1.4 p-value

On se donne une famille de règles de décisions $d_\alpha : \mathcal{X} \to \{0; 1\}$ pour tout niveau α dans $[0; 1]$. On note $\mathcal{R}_\alpha = \{x \in \mathcal{X} \ : \ d_\alpha(x) = 1\}$ les zones de rejet associées. On fait les hypothèses suivantes :

1. pour tout α,
$$\sup_{\theta\in\mathcal{T}_0} \mathbb{P}_\theta\left[X \in \mathcal{R}_\alpha\right] = \alpha \ ; \tag{2}$$
2. pour tout $0 \le \alpha \le \alpha' \le 1$, on a $\mathcal{R}_\alpha \subset \mathcal{R}_{\alpha'}$.

Remarque 6. L'hypothèse 2 est équivalente à dire que, pour tout $x \in \mathcal{X}$, $\alpha \mapsto d_\alpha(x)$ est croissante.

Définition 7. La p-value est la plus petite valeur de α pour laquelle on rejette H_0 sur les données observées. Formellement, pour tout $x \in \mathcal{X}$,
$$p(x) = \inf\{\alpha \ : \ x \in \mathcal{R}_\alpha\}.$$

Remarque 8.

— La p-value est une statistique. C'est une fonction du jeu de données que l'on peut calculer sans connaître la valeur du paramètre θ inconnue.
— La p-value **n'**est **pas** la probabilité de l'hypothèse alternative !
— C'est la valeur de α pour laquelle la décision change sur les données.

Proposition 9. *On suppose que le suprémum de l'équation* (2) *est atteint en un unique θ_0 qui ne dépend pas de α. Si $X \sim P_{\theta_0}$, alors la statistique $p(X)$ de p-value est distribuée suivant la loi uniforme sur l'intervalle* $[0\ ;\ 1]$.

Exercices 1, 2, 3, 4.

2 Test sur μ et σ^2 dans un échantillon gaussien.

On suppose qu'on observe un n-échantillon $(X_1, \cdots, X_n)$ de la loi $\mathcal{N}(\mu, \sigma^2)$. On considère les estimateurs empiriques de μ et σ^2 :
$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i\,, \qquad S_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2. \tag{3}$$
et l'estimateur sans biais de σ^2 :
$$S_n'^2 = \frac{n}{n-1} S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

Définition 10. Loi du χ^2 et de Student.

1. Soit $(Y_1, \cdots Y_p)$ un p-échantillon de loi $\mathcal{N}(0,1)$. La variable $Z = \sum_{i=1}^p Y_i^2$ suit **une loi du χ^2 à p degrés de liberté (χ^2_p)**. La densité de Z est donnée par
$$\frac{1}{2^{p/2}\Gamma(p/2)} e^{-x/2} x^{p/2-1} \ \mathbb{1}_{x\ge 0}\,.$$
2. Soit Y et Z deux variables indépendantes, $Y \sim \mathcal{N}(0,1)$, $Z \sim \chi^2_p$. On appelle **loi de Student à p degrés de liberté (St_p)**, la loi de la variable $Y/\sqrt{Z/p}$.

.

Proposition 11. Loi de ces estimateurs.
1. *$\bar{X}_n$ est de loi $\mathcal{N}(\mu, \frac{\sigma^2}{n})$.*
2. *$(nS_n^2)/\sigma^2$ est de loi χ^2_{n-1}.*
3. *Les variables $\bar{X}_n$ et S_n^2 sont indépendantes.*

Proposition 12. Biais, consistance.
1. *Les estimateurs $\bar{X}_n$ et S_n^2 sont consistants.*
2. *$\bar{X}_n$ est un estimateur sans biais et efficace de μ.*
3. *$S_n'^2$ est un estimateur sans biais de σ^2.*

Proposition 13. Test sur μ avec σ inconnue.
Soit $t_{n-1,\alpha}$ le quantile d'ordre α de St_{n-1}. L'ensemble $\mathcal{R} = \left\{\frac{|\bar{X}_n - \mu_0|}{S_n'^2/\sqrt{n}} \ge t_{n-1;\alpha/2}\right\}$ est la région de rejet d'un test de niveau α de (H_0) : "$\mu = \mu_0$" contre (H_1) : "$\mu \neq \mu_0$" .

De la même façon, $\mathscr{R} = \left\{\bar{X}_n < \mu_0 + t_{n-1,\alpha}\frac{S'_n}{\sqrt{n}}\right\}$, est la région de rejet d'un test de niveau α de (H_0) : "$\mu \geq \mu_0$" contre l'hypothèse (H_1) : "$\mu < \mu_0$".

Proposition 14. test sur σ^2.
Soit $c_{n-1,\alpha}$ le quantile d'ordre α de χ^2_{n-1}. L'ensemble $\mathscr{R} = \left\{(n-1)\frac{S'^2_n}{\sigma_0^2} \leq c_{n-1,\alpha}\right\}$ est la région de rejet d'un test de niveau α de (H_0) : "$\sigma \geq \sigma_0$" contre l'hypothèse (H_1) : "$\sigma < \sigma_0$".

Exercices 5, 6.

3 Deux types de statistiques de tests

3.1 Test de Wald

Le test de Wald suppose que l'on dispose d'un estimateur $\hat{\theta}_n$ de θ **asymptotiquement normal**, i.e. tel que

$$\frac{\hat{\theta}_n - \theta}{\hat{\sigma}_n} \xrightarrow[n\to+\infty]{\text{loi sous } \mathbb{P}_\theta} \mathcal{N}(0,1),$$

où $\hat{\sigma}_n$ est un estimateur de l'écart-type de $\hat{\theta}_n$.

Supposons par exemple que l'on veuille tester (H_0) : "$\theta \leq \theta_0$", contre (H_1) : "$\theta > \theta_0$", où θ_0 est une valeur fixée. On peut prendre la règle de décision :
- Si $\frac{\hat{\theta}_n - \theta_0}{\hat{\sigma}_n} > t$, on rejette (H_0) ;
- Si $\frac{\hat{\theta}_n - \theta_0}{\hat{\sigma}_n} \leq t$, on ne rejette pas (H_0).

t est à choisir tel que

$$\alpha = \mathbb{P}_{(H_0)}(\text{Rejeter } (H_0)) = \sup_{\theta\leq\theta_0} \mathbb{P}_\theta\left[\frac{\hat{\theta}_n - \theta_0}{\hat{\sigma}_n} > t\right] = \sup_{\theta\leq\theta_0} \mathbb{P}_\theta\left[\frac{\hat{\theta}_n - \theta}{\hat{\sigma}_n} + \frac{\theta-\theta_0}{\hat{\sigma}_n} > t\right] \approx \sup_{\theta\leq\theta_0} \mathbb{P}\left[Z + \frac{\theta-\theta_0}{\hat{\sigma}_n} > t\right]$$

où $Z \sim \mathcal{N}(0,1)$. La fonction $\theta \mapsto \mathbb{P}\left[Z + \frac{\theta-\theta_0}{\hat{\sigma}_n} > t\right]$ est croissante en θ, et le sup est atteint en θ_0. On a donc

$$\alpha \approx \mathbb{P}[Z > t],$$

et t est le quantile d'ordre $1-\alpha$ de la loi $\mathcal{N}(0,1)$.

3.2 Test du rapport de vraisemblance

Supposons que l'on veuille tester (H_0) : "$\theta \in \mathscr{T}_0$", contre (H_1) : "$\theta \in \mathscr{T}_1$", où $\mathscr{T}_0 \cup \mathscr{T}_1$ est une partition de $\mathscr{T}$.

Il est raisonnable de rejeter (H_0) si l'observation faite X est beaucoup plus probable quand le paramètre θ varie dans $\mathscr{T}_1$, que lorsqu'il varie dans $\mathscr{T}_0$, i.e si

$$\sup_{\theta\in\mathscr{T}_1} L(\theta, X) \gg \sup_{\theta\in\mathscr{T}_0} L(\theta, X)$$

Ainsi, le test du rapport de vraisemblance s'appuie sur la statistique $T(X) = 2\log\frac{\sup_{\theta\in\mathscr{T}_1} L(\theta,X)}{\sup_{\theta\in\mathscr{T}_0} L(\theta,X)}$ (T ne dépend que de X et pas du paramètre θ). Une région de rejet raisonnable est donc

$$\mathscr{R} = \{X \text{ tel que } T(X) \geq t\}$$

où t est à choisir en fonction du niveau du test.

Pour choisir t, on doit donc connaître la loi de $T(X)$ sous (H_0), ou au moins sa loi asymptotique quand n tend vers l'infini.

Lorsqu'on teste (H_0) : "$\theta = \theta_0$" contre (H_1) : "$\theta \neq \theta_0$", on peut remarquer que dès que $L(\theta, X)$ est continue en θ,

$$\sup_{\theta\in\mathscr{T}_0} L(\theta, X) = L(\theta_0, X),$$

$$\sup_{\theta\in\mathscr{T}_1} L(\theta, X) = \sup_{\theta\neq\theta_0} L(\theta, X) = \sup_{\theta\in\mathscr{T}} L(\theta, X) = L(\hat{\theta}, X),$$

où $\hat{\theta}$ est l'estimateur du maximum de vraisemblance. Ainsi

$$T(X) = 2(\ell(\hat{\theta}, X) - \ell(\theta_0, X)).$$

On a le résultat suivant :

Proposition 15. *$2(\ell(\hat{\theta}, X) - \ell(\theta, X))$ converge en loi (sous $\mathbb{P}_\theta$) vers la loi du χ^2_d, où d est la dimension de l'espace des paramètres.*

Idée de preuve : On fait un développement de Taylor de ℓ autour de $\hat{\theta}$:

$$\begin{aligned}\ell(\theta, X) &= \ell(\hat{\theta}, X) + \nabla\ell(\hat{\theta}, X)(\theta - \hat{\theta}) + \frac{1}{2}\langle\theta - \hat{\theta}, \frac{\partial^2}{\partial\theta^2}\ell(\theta, X)(\theta - \hat{\theta})\rangle + \text{ reste} \\ &= \ell(\hat{\theta}, X) + \frac{1}{2}\langle\theta - \hat{\theta}, \frac{\partial^2 \ell}{\partial\theta^2}(\hat{\theta}, X)(\theta - \hat{\theta})\rangle + \text{ reste}\end{aligned}$$

On utilise ensuite la loi asymptotique de l'EMV vue au chapitre 2 : $I_n^{1/2}(\theta)(\hat{\theta} - \theta) \overset{\text{loi}}{\to} \mathcal{N}_d(0, \mathrm{Id})$ quand $n \to +\infty$, pour conclure. □

Ainsi, sous (H_0) : "$\theta = \theta_0$", $T(X)$ converge en loi vers la loi du χ^2_d. On peut donc terminer la construction du test, et déterminer la valeur de t :

$$\alpha = \mathbb{P}_{(H_0)}(\text{Rejeter } (H_0)) = \mathbb{P}_{\theta_0}(T(X) \geq t) \approx \mathbb{P}(Z \geq t), \text{ où } Z \sim \chi^2_d.$$

t est donc le quantile d'ordre $1 - \alpha$ de la loi du χ^2_d.

Lorsqu'on teste (H_0) : "$\theta = \theta_0$" contre (H_1) : "$\theta = \theta_1$", où θ_0 et θ_1 sont deux valeurs fixées de θ, $T(X) = 2(\log(L(\theta_1, X)) - \log(L(\theta_0, X))$. Supposons qu'il existe une unique valeur k_α telle que $\mathbb{P}_{\theta_0}(T(X) \geq k_\alpha) = \alpha$, de telle sorte que $\mathscr{R}_0 = \{X; T(X) \geq k_\alpha\}$ est la région de rejet du test du rapport de vraisemblance de niveau α.

On a alors

Lemme 16. Lemme de Neyman-Pearson.
Soir $\mathscr{R}$ la région de rejet d'un test de niveau α de (H_0) : "$\theta = \theta_0$" contre (H_1) : "$\theta = \theta_1$". Alors $\mathbb{P}_{\theta_1}(\mathscr{R}_0) \geq \mathbb{P}_{\theta_1}(\mathscr{R})$. Autrement dit, le test du rapport de vraisemblance est plus puissant que tout autre test de niveau α.

Idée de preuve : On vérifie que par définition de $\mathscr{R}_0$,

$$(\mathbb{1}_{\mathscr{R}_0} - \mathbb{1}_{\mathscr{R}})(L(\theta_1, X) - e^{k_\alpha/2}L(\theta_0, X)) \geq 0.$$

Ainsi,

$$\begin{aligned}\mathbb{P}_{\theta_1}(\mathscr{R}_0) - \mathbb{P}_{\theta_1}(\mathscr{R}) &= \int (\mathbb{1}_{\mathscr{R}_0}(x) - \mathbb{1}_{\mathscr{R}}(x))L(\theta_1, x)\, d\mu(x) \\ &\geq e^{k_\alpha/2}\int (\mathbb{1}_{\mathscr{R}_0}(x) - \mathbb{1}_{\mathscr{R}}(x))L(\theta_0, x)\, d\mu(x) = e^{k_\alpha/2}(\mathbb{P}_{\theta_0}(\mathscr{R}_0) - \mathbb{P}_{\theta_0}(\mathscr{R})) = 0\end{aligned}$$

□

Exercice 4.

4 Exercices.

Exercice 1. On rappelle qu'une région de confiance $I_n(\omega)$ de coefficient de sécurité $1 - \alpha$ est définie par :

$$\mathbb{P}_\theta\,[\theta \in I_n(\omega)] \geq 1 - \alpha, \quad \forall\, \theta \in \Theta.$$

Montrer que pour toute région de confiance $I_n(\omega)$, il existe un test de H_o : "$\theta = \theta_o$" contre H_1 : "$\theta \neq \theta_o$" , tel que

$$I_n(\omega) = \{\theta_o \mid [\text{on accepte } H_o(\theta_o)](\omega)\}.$$

Exercice 2. Soit X un échantillon de taille n de la loi $N(\theta; 1)$. On sait que $\theta \in \Theta = \{\theta_0; \theta_1\}$. $\theta_0 < \theta_1$. On veut tester (H_0) : " $\theta = \theta_0$" contre (H_1) : " $\theta = \theta_1$". On considère le test de région de rejet $R(X) = \left\{\bar{X} \geq \frac{\theta_0 + \theta_1}{2}\right\}$. Calculer le niveau du test. Soit α fixé, quelle valeur de n faut-il prendre pour que le test soit de niveau α? Calculer la puissance du test.

Exercice 3. Soit X un échantillon de taille n de la loi de Poisson de paramètre θ, $\theta \in \Theta = \{1;2\}$. On considère le test de (H_0) : " $\theta = 1$" contre (H_1) : " $\theta = 2$" de région de rejet $R(X) = \{\bar{X} > 3\}$. Si on veut un test de niveau 5%, comment faut-il choisir n ? Calculer alors la puissance du test.

Exercice 4. Soit $X = (X_1, \cdots, X_n)$ un échantillon de taille n de la loi uniforme sur $[0;\theta]$. On note $\hat{\theta}_n = \max(X_1, \cdots, X_n)$.

1. Montrez que $\hat{\theta}_n/\theta$ est un pivot pour l'estimation de θ.
2. Construire un test de (H_0) : " $\theta = 1$" contre (H_1) : " $\theta \neq 1$" de niveau α.
3. Quelle est la courbe de puissance de ce test ?
4. Le test construit est-il le test du rapport de vraisemblance ?
5. Construire le test de Wald basé sur $\hat{\theta}_n$.

Exercice 5. Un produit commercialisé est présenté dans des boites sur lesquelles on peut lire : *contenance 500 grammes*. On peut se demander ce que signifie cette donnée car la quantité contenue dans une boite choisie au hasard est une variable aléatoire supposée à densité (et désignée par X). Soit μ la moyenne de la variable X. Les informations recueillies auprès du fabricant permettent d'affirmer que dans ce cas la valeur de μ est censée être égale à 500 grammes.

1. Donner un intervalle de confiance pour μ de coefficient de sécurité 95%.
2. Tester si la norme du fabricant est respectée.

Données (en grammes) : 490 ; 490 ; 490 ; 492 ; 492 ; 495 ; 497 ; 497 ; 502 ; 505.

Exercice 6. Une usine fabrique des câbles dont la charge de rupture suit une loi $\mathcal{N}(\mu_o, \sigma^2)$ avec $\mu_o = 99$ kg. Pour tester un nouveau procédé de fabrication des câbles, on a observé sur dix câbles les charges de ruptures suivantes (en kilos) : 101 ; 102 ; 100 ; 104 ; 105 ; 99 ; 103 ; 100 ; 101 ; 105 . Le nouveau procédé est-il meilleur que le précédent ? Sur l'ancien procédé de fabrication, on avait observé $\sigma = 1kg$. Le nouveau procédé est-il plus précis que l'ancien ?